

Large deformation 3D image registration in image-guided radiation therapy

Mark Foskey, Brad Davis, Lav Goyal, Sha Chang, Ed Chaney, Nathalie Strehl, Sandrine Tomei, Julian Rosenman and Sarang Joshi

Department of Radiation Oncology, University of North Carolina

E-mail: mark_foskey@unc.edu

Abstract. In this paper, we present and validate a framework, based on deformable image registration, for automatic processing of serial 3D CT images used in image-guided radiation therapy. A major assumption in deformable image registration has been that, if two images are being registered, every point of one image corresponds appropriately to some point in the other. For intra-treatment images of the prostate, however, this assumption is violated by the variable presence of bowel gas. The framework presented here explicitly extends previous deformable image registration algorithms to accommodate such regions in the image for which no correspondence exists.

We show how to use our registration technique as a tool for organ segmentation, and present a statistical analysis of this segmentation method, validating it by comparison with multiple human raters. We also show how the deformable registration technique can be used to determine the dosimetric effect of a given plan in the presence of non-rigid tissue motion. In addition to dose accumulation, we describe a method for estimating the biological effects of tissue motion using a linear-quadratic model. This work is described in the context of a prostate treatment protocol, but it is of general applicability.

Submitted to: *Phys. Med. Biol.*

1. Introduction

In radiation cancer therapy, the problem of organ motion over the course of treatment is becoming more urgent as techniques for conformal therapy improve. These techniques, such as intensity modulated radiation therapy (IMRT), offer important benefits: With high gradients between the region receiving a therapeutic dose and surrounding regions, it is possible, in principle, to increase the prescribed dose to the tumor while reducing the dose to critical organs. The problem with these high gradients is that organ location varies between treatment days, because of both setup error and internal changes such as bowel and bladder filling. With high dose gradients, relatively little organ motion is required to bring parts of the tumor outside of the therapeutic region, or to bring healthy critical tissues in. Both forms of tissue misplacement can harm the patient, in the one case by failure of local control, and in the other, by toxicity to normal tissue. There are now in-the-treatment-room imaging methods, such as cone beam CT and CT-on-rails, that enable image guided radiation therapy as a way to meet this challenge. However, there remains a pressing need for automatic techniques to translate these images into useful information about organ location and likely treatment effectiveness.

The traditional approach to the problem of organ motion has been to specify a margin around the clinical target volume (CTV) to create the planning target volume (PTV). The goal of the margin is to achieve a specified confidence level, interpreted as the probability, at a given treatment session, that actual tumor is contained entirely within the PTV. Work by Goitein and Busse (1975) and Goitein (1985, 1986) suggests that a confidence level of 95% is required. Typically, the size of the margin is expressed as a single parameter, its width, which is based on studies of organ motion across populations of patients. Sometimes the width is reduced near critical structures. For instance, with prostate cancer, the size of the margin may be set to 1 cm, with a reduction to 6 mm toward the rectum (Happersett *et al* 2003).

This simple construction of the PTV relies on two assumptions that have been necessitated by technical limitations in treatment planning and delivery. The first assumption is that organ motion has the same statistical properties for different patients, so that the variance in organ position for a single patient will be equal to that computed previously for a population of patients. The second assumption is that organ motion is statistically the same for all parts of the organ.

To avoid having to make the first assumption, Yan *et al* (1997) introduced the framework of adaptive radiation therapy (ART), in which organ motion for the individual patient is measured over the course of treatment, and the PTV is modified once the amount of motion for that patient has been estimated with sufficient confidence. In their work, the position variation is expressed as a single parameter, a 95% confidence radius for the position of the tumor isocenter, thus still making the assumption that motion is uniform across the relevant organs.

To account for motion that is not uniform, in which organs deform and move relative to one another, a more sophisticated analysis of images is necessary. Recent

computational advances have enabled the emergence of a discipline called computational anatomy (Grenander and Miller 1998) with the principal aim of developing specialized mathematical and software tools for the precise mathematical study of anatomical variability. Within computational anatomy, *deformable image registration* techniques have proved to be effective in the study of anatomical variation (Davatzikos 1996, Christensen *et al* 1997, Csernansky *et al* 1998, Joshi *et al* 1997, Thompson and Toga 2002).

In the framework of computational anatomy, this paper presents a comprehensive approach for automatic processing of 3D CT images acquired during image guided radiation therapy. Deformable image registration is the key to the approach, making it possible to establish a correspondence between points in images taken on different days. Such a correspondence is useful in two key ways: It facilitates automatic organ segmentation, and it makes it possible to calculate the dosimetric effects of nonrigid tissue motion.

The need for careful repeated segmentations has been one of the major limitations for the widespread application of ART and other image guided techniques. Although careful manual segmentation techniques remain the standard of practice, a full manual segmentation of the intra-treatment CT images is time consuming, expensive and not practical in a routine clinical setting. Moreover, manual segmentation introduces uncertainties associated with variability both between and within raters. Two European studies that focused on user-guided tumor segmentation found large inter-user variabilities for well circumscribed lesions (Leunens *et al* 1993, Valley and Mirimanoff 1993).

The dosimetric analysis of tissue motion has the potential to permit more sophisticated ART planning than is currently being pursued (Birkner *et al* 2003). A number of groups have studied the dosimetry of rigid patient motion (Booth and Zavgorodni 2001, Booth 2002, Unkelbach and Oelfke 2004), and there has also been some work in dosimetric analysis of deforming tissue (Schaly *et al* 2004, Yan *et al* 1999). The registration algorithm we describe here differs from previous work in that it provides a fully automated means of performing dose accumulation that can handle large deformations.

In the context of radiotherapy of the prostate or cervix, several deformable image registration methods are currently being investigated for alignment of serial CT data sets. Schaly *et al* (2004) use an approach based on thin-plate splines (Bookstein 1989) for matching CT volumes, where homologous points are chosen from manually drawn organ segmentations. They use the resulting displacement fields to measure cumulative dose over multiple fractions for prostate cancer patients. Christensen *et al* (2001) reported registration of serial CT images for patients undergoing treatment for cervix cancer. Their method matches the boundaries of the bladder, rectum, and vagina/uterus, which are first manually segmented in the planning and treatment images. As with our work, they use a viscous-fluid model that accommodates large deformation changes in the anatomy. Wang *et al* (2005) register CT volumes using a method similar

to the demons algorithm of Thirion (1998). Their method employs a voxel-based driving force motivated by optical flow and a Gaussian regularization kernel. They provide an example of automatic segmentation of a treatment image using the resulting deformation fields. Lu *et al* (2004) present a deformable registration technique based on the minimization of an energy functional that combines an image matching term with a smoothness measure on the resulting deformation field. However, none of these studies address the problem of bowel gas for deformable registration of CT images. Also, while some authors have presented validation studies based on known transformations or phantoms, to our knowledge none have presented a large scale analysis of the accuracy of their methods for automatic segmentation of treatment images based on manual contours.

To give background for what follows, we briefly describe the ART protocol (adapted from Yan *et al* 2000) that we use in our regular prostate care. The fundamental purpose is to use a planning target volume (PTV) that reflects the typical organ motion of the particular patient. Rather than attempting to determine that motion prior to treatment, we use a conventional plan during the first five treatment days, at the same time acquiring a registered CT scan each day. After the fifth treatment day, we construct a new PTV by placing a margin around the approximate convex hull of the CTVs from the first five treatment days, and then generate a new plan, this time using IMRT, based on the new PTV. For the remainder of the treatment period, images are acquired twice weekly to indicate whether further adjustments may be necessary. For each image, the patient is first set up for treatment using crosshair tattoos that are aligned with laser fiducials. Then CT-visible skin markers (2.3-mm “BBs”) are placed at the locations marked by the lasers, so that the treated isocenter is indicated on the scan. In a future paper we will assess the effectiveness of this protocol in our practice, using the dosimetric techniques described in this paper.

Shown in figure 1 is a visualization of the organ motion over the course of treatment for 9 patients treated in our clinic using the ART protocol. The internal organ motion of the prostate shown in the images was estimated using manual segmentations of intra-treatment CT images acquired by the CT-on-rails system.

The rest of the paper will be organized as follows. In Section 2 we explain the registration algorithms that we use. In Section 3 we explain how we use deformable registration as a tool for segmentation, and evaluate the reliability of the resulting segmentations. In Section 4 we explain dosimetric applications of our algorithms, and we conclude in Section 5.

2. Deformable image registration

The key to our approach is the measurement of organ motion by means of deformable image registration. We interpret the term “organ motion” broadly, to include setup error and any internal tissue displacement or deformation. We measure organ motion by comparing a CT image taken at planning time to a *treatment image* taken immediately

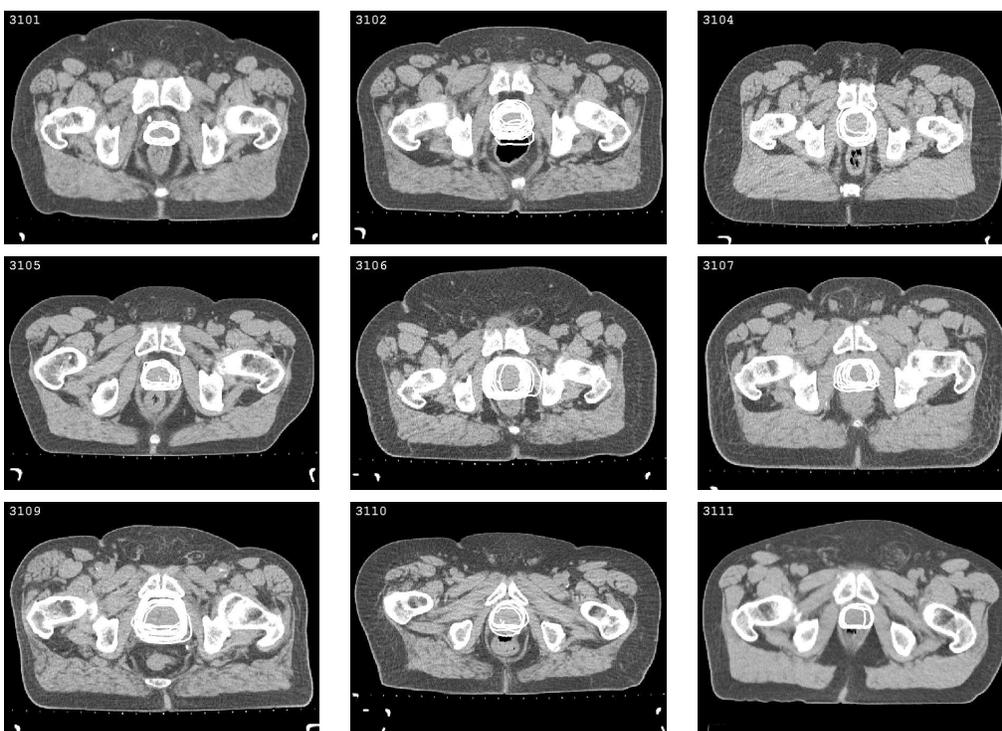


Figure 1. Visualization of prostate motion over the course of treatment for 9 patients involved in our study. White contours, superimposed on an axial slice of each patient’s planning image, indicate the actual location of the prostate on each treatment day. These contours are taken from manual segmentations of treatment images. The discrepancies between the contours exhibit the effect of setup error and organ motion on the prostate position. Note that different patients exhibit different amounts of prostate motion; compare the close contour agreement for patient 3101 with the wide contour variability for patient 3109. For some patients (3102, 3109) motion is primarily noticeable in the anterior-posterior direction; for other patients (3106, 3107) motion is primarily noticeable in the lateral direction.

before a given treatment, both of which are acquired using a Siemens Primatom system that provides a CT scanner sharing a table with the treatment machine. If there were no organ motion, the planning image and all the treatment images would be the same, except for noise from the imaging device. However, because there is organ motion, these images will differ, and the difference characterizes the motion (figure 2).

Figure 3 compares a difference image between two unregistered images (aligned as treated) to the difference image for the same two images after registration has been performed. We have understood the motion when we can tell, for each point in the planning image, which point in the treatment image it corresponds to. In this way organ motion and image registration are linked—we can understand organ motion if we can estimate image correspondence. Once image correspondence is established, contours of structures such as the tumor body can be transformed, and other detailed analysis of the changes can be done. The purpose of this section is to explain the registration algorithms we use to establish the correspondence.

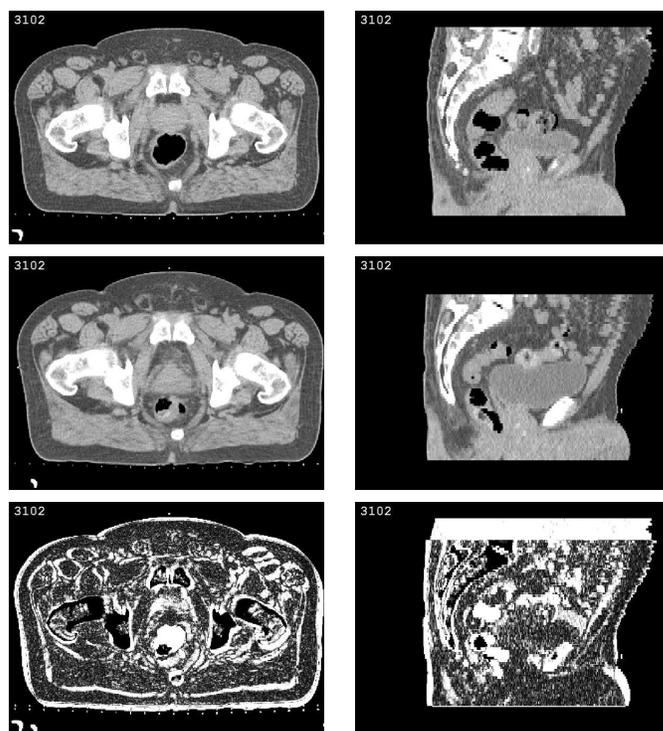
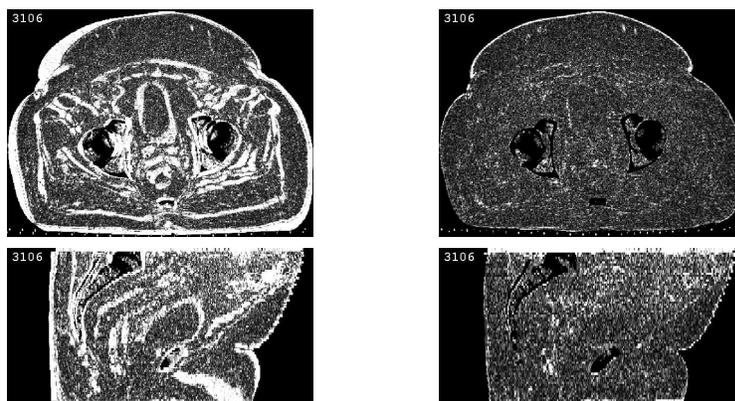


Figure 2. First row: axial and sagittal slices from the planning image of patient 3102. Second row: the same slices (with respect to the planning image coordinate system) taken from a treatment image. Third row: the voxelwise absolute difference between the planning and treatment images. Black represents perfect intensity agreement, which is noticeable in the interior of the bones and outside the patient. Brighter regions, indicating intensity disagreement, are especially apparent: (1) in regions where gas is present in one image and absent in the other, (2) around the bladder which is large on the treatment day compared to the planning day, (3) uniformly along boundaries with high intensity gradient, indicating a global setup error such as a translation.



Difference Before Registration Difference After Registration

Figure 3. Difference images comparing a planning to a daily image before and after deformable registration.

We use the term *tissue voxel* to refer to a volume of tissue small enough to be considered as a single point for the purposes of analysis. We view an image as a function $I(x)$ from a domain $V \subset \mathbb{R}^3$ to \mathbb{R} , so that $I(x)$ is the intensity of the image at the point $x \in V$. Then the image correspondence can be expressed as a function $h: V \rightarrow V$, called a *deformation field*. For $x \in V$, $h(x)$ is the point in the treatment image that corresponds to x in the planning image. To the extent that the image registration corresponds to the tissue motion, $h(x)$ is the location, at treatment time, of the tissue voxel originally at x . We find $h(x)$ by approximately minimizing an energy term

$$E(h) = \int_V \left(I_P(x) - I_T(h(x)) \right)^2 dx, \quad (1)$$

subject to an appropriate regularity condition. It makes sense to minimize the squared differences of image intensities directly because the CT intensities (expressed in Hounsfield units) have direct physical meaning. The fact that the same machine is used to acquire all images reduces the chance of calibration error.

We decompose the motion into two components, a global rigid transformation (translation and rotation) followed by a deformation that allows the soft tissue to align. This decomposition improves performance since the rigid alignment is fast and accounts for a large portion of the image misalignment. It also makes sense from a clinical perspective since the rigid misalignment corresponds closely to patient setup error and can thus be used to provide guidance for improving setup techniques.

There is one point about image registration that is worth emphasizing. In our formulation, h maps the space of I_P to that of I_T . But we use it to deform I_T , by composing I_T with h , creating a new image that we could write $I_T^{\text{deformed}}(x) = I_T(h(x))$. This approach makes it straightforward to calculate the new image: For each voxel x , we use $h(x)$ to look up an intensity in I_T , interpolating if necessary. If we wish to deform I_P , we compute h^{-1} and then evaluate $I_P(h^{-1}(x))$.

2.1. Rigid Motion

We have used both translation and general rigid motion in our work. For clarity we will explain our algorithm for translation first, after which we will place this method in a more general setting, which we use to perform general affine and rigid registration.

In the case of translation, we want to minimize the energy E subject to the condition that $h(x)$ is of the form $x + \tau$ for some translation vector τ . Thus (1) becomes

$$E(\tau) = \int_V \left(I_P(x) - I_T(x + \tau) \right)^2 dx$$

Following Joshi *et al* (2003), we use a quasi-Newton algorithm to minimize $E(\tau)$, constructing a sequence $\{\tau_k\}$ such that $E(\tau_k)$ converges to a local minimum. Let $\tau_{k+1} = \tau_k + \Delta\tau_k$; we will derive a formula for $\Delta\tau_k$. For convenience, write $x' = x + \tau_k$. If we expand $I_T(x + \tau_{k+1}) = I_T(x' + \Delta\tau_k)$ in a first order Taylor series about x' , we get

$$E(\tau_{k+1}) \approx \int_V \left(I_P(x) - I_T(x') + \nabla I_T(x') \cdot \Delta\tau_k \right)^2 dx.$$

At each step in the iteration we find the $\Delta\tau_k$ that minimizes our approximation to $E(\tau_{k+1})$, by setting its gradient to 0 and solving. We get

$$\Delta\tau_k = \left(\int_V \nabla I_T(x') \nabla I_T(x')^T dx \right)^{-1} \int_V \left(I_P(x) - I_T(x') \right) \nabla I_T(x') dx. \quad (2)$$

In a more general setting, we consider a transformation h that depends on a parameter vector a as well as x , so that we may write $h = h_a(x)$. We then want to find Δa_k . The expression $I_T(h_a(x))$ is a function of both x and a , and, in the same way that (2) was derived, we find that

$$\begin{aligned} \Delta a_k = & \left(\int_V \nabla_a I_T(h_a(x)) \nabla_a I_T(h_a(x))^T dx \right)^{-1} \\ & \times \int_V \left(I_P(x) - I_T(h_a(x)) \right) \nabla_a I_T(h_a(x)) dx. \end{aligned} \quad (3)$$

An important example is the case of h an affine transformation, that is, of the form $h(x) = Ax + \tau$ for some matrix A and translation vector τ . In this situation, $\nabla_a I_T(h_a(x))$ can be expressed conveniently in the following way. We define the parameter vector a by

$$a = [A_{11} \ A_{12} \ A_{13} \ A_{21} \ \dots \ A_{32} \ A_{33} \ \tau_1 \ \tau_2 \ \tau_3]^T.$$

We then define, for any point $x = (x_1, x_2, x_3)$,

$$X = \begin{bmatrix} x_1 & x_2 & x_3 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & x_1 & x_2 & x_3 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & x_1 & x_2 & x_3 & 0 & 0 & 1 \end{bmatrix},$$

so that $Ax + \tau = Xa$. With this convention, $\nabla_a I_T(h_a(x)) = (\nabla I_T|_{h_a(x)})^T X$, which can be easily computed and used in (3). Here we use the notation $\nabla I_T|_{h_a(x)}$, rather than $\nabla I_T(h_a(x))$, to indicate that the gradient $\nabla I_T(\cdot)$ is calculated first, with the result simply evaluated at $h_a(x)$.

For *rigid* registration, as opposed to affine, at each iteration we perform the the same step as for affine registration, and then replace the resulting matrix A by the rotation matrix that most closely approximates it. To find the rotation matrix we calculate the *polar decomposition* $A = RD$ (Horn and Johnson 1990), where R is an orthogonal matrix and D is positive semidefinite, and take the matrix R as our approximation to A . This decomposition is unique provided that A is invertible, which holds in practice for medical images.

2.2. Deformation

In the case of large deformation registration, rather than constraining h by requiring that it be expressed in a specific form, we modify the energy functional by adding a regularity term that quantifies how severely h deforms the image. Thus we get

$$E(h) = \int_V \left(I_P(x) - I_T(h(x, t)) \right)^2 dx + E_{\text{reg}}(h).$$

In Bayesian terms, the first term is a likelihood estimate, and the second is a kind of prior on the space of transformations. The key difficulty in this kind of registration is to find a prior that permits large deformations but not arbitrary rearrangements of voxels. The solution that we adopt was first detailed by Christensen *et al* (1996) and further developed by Joshi and Miller (2000). The idea is to introduce a time parameter t and define a function $h(x, t)$ such that $h(x, 0) = x$ and $h(x, t_{\text{final}})$ is the desired deformation field $h(x)$ that aligns I_P and I_T . We construct h as the integral of a time-varying velocity field:

$$h(x, t) = x + \int_0^t v(h(x, s), s) ds,$$

and we define

$$E_{\text{reg}}(h) = \int_{V,t} \|L_{\text{reg}}v(x, t)\|^2 dx dt$$

where L_{reg} is some suitable differential operator. In this way, the size of E_{reg} is not directly based on the difference between $h(x)$ and x , which would tend to prevent large deformations. In the context of landmark-based image registration, Joshi and Miller (2000) show that this method, with proper conditions on L_{reg} , produces a diffeomorphism (i.e. differentiable with a differentiable inverse). As a result, each position x in the planning image corresponds to a unique position in the treatment image, and no tearing of tissue occurs.

Optimization of the resulting functional $E(h)$ is computationally intensive, since the velocity vector fields for all time steps must be optimized together (Miller *et al* 2002, Beg *et al* 2005). Therefore we follow a greedy approach. At each time step, we choose the velocity field that improves the image match most rapidly, subject to the smoothness prior. Precisely, for each t we minimize

$$\frac{d}{ds} \int_V \left(I_P(x) - I_T(h(x, t) + sv(h(x, t), t)) \right)^2 dx \Big|_{s=0} + \int_V \|L_{\text{reg}}v(x, t)\|^2 dx.$$

After evaluating the derivative and solving the resulting variational problem, we find that v must satisfy the differential equation

$$(I_P(x) - I_T(h(x, t))) \nabla I_T(h(x, t)) = Lv(x, t), \quad (4)$$

where L is a differential operator proportional to $(L_{\text{reg}})^\dagger L_{\text{reg}}$.

A number of choices of L are reasonable, depending on the desired behavior of the algorithm. We choose the operator $Lv = \alpha \nabla^2 v + \beta \nabla(\nabla \cdot v) + \gamma v$, a choice motivated by the Navier-Stokes equations for compressible fluid flow with negligible inertia. Note that the Laplace operator ∇^2 is applied to each component of v separately.

If we interpret v as the velocity field of a fluid, then the left hand side of (4) represents an image force exerted on each point in the fluid domain. Note that, at each point, the force is along the direction of greatest change in image intensity of $I_T(h(x, t))$, and the magnitude and sign of the force is determined by the difference in intensity between the two images. The right hand side of the equation expresses the resistance to flow. This notional fluid has the nonphysical property that it resists

compression (and dilation) inelastically, so that volume can be permanently added or removed in response to image forces. Also, the γ term, which can be thought of as a “body friction” term, ensures that L is a positive definite differential operator, and hence invertible (Joshi and Miller 2000).

To compute $h(x)$, we integrate the resulting velocity field forward in time until the change in image match between successive time steps drops below a threshold. At each time step we find v , using the fast Fourier transform, by explicitly inverting L in the frequency domain. In the context of landmark-based image registration, Joshi and Miller (2000) show that this method, with proper conditions on L , does produce a diffeomorphism. To make sure that Euler integration, being discrete, does not introduce singularities, we choose a step size such that the largest distance moved by a voxel between successive time steps is less than the inter-voxel spacing.

2.3. Bowel Gas

In images of the pelvic region, one problem that arises in deformable image registration is associated to the presence of bowel gas. Regions of gas appear as black blobs surrounded by gray tissue (see figures 2 and 4). Typically, there will not be gas at the same location in the intestine for different images, and in that case there is no reasonable diffeomorphism between the domains of the two images. That is, if $x \in V$ is in a region containing gas in the planning image, and there is no intestinal gas in the same part of the treatment image, then there is no location in the treatment image that naturally corresponds to x , and thus no reasonable value for $h(x)$. Solid bowel contents do not produce the same difficulty because they do not contrast greatly with the inner wall of the bowel, and are therefore handled by the compressibility of the fluid flow model.

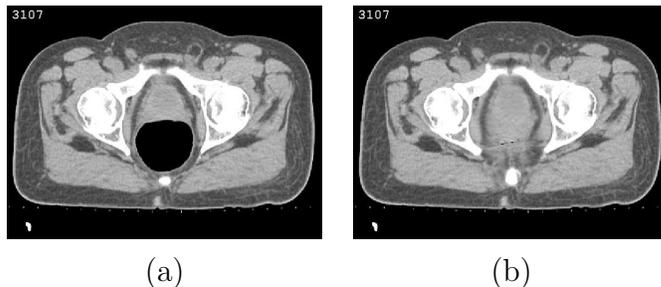


Figure 4. Example of gas deflation. Panel (a) shows an axial slice of a treatment image containing a large region of bowel gas. Panel (b) shows the same image after automatic gas deflation. This deflated image can be accurately registered using deformable image registration.

To resolve the problem of gas, we process each image exhibiting the problem to shrink the gassy region to a point, using a variation of our image deformation algorithm that we refer to as *deflation*. This algorithm is not meant to simulate the true motion of the tissue but to eliminate the gas in a principled way so that the image can be

accurately registered. Deflation needs to be applied to both the planning and treatment images, if they both have gas, since the pockets are typically in different places. In this section, we describe the deflation algorithm itself, and in the next, we explain how we combine deflation with the deformable and translation registration algorithms to establish the correspondence between the planning and treatment images.

The algorithm is defined as follows. We first threshold the image so that gas appears black and tissue appears white, which is possible since the contrast between gas and surrounding tissue is very high in CT images. We refine this binary segmentation by a morphological opening, eroding and then dilating the gas region, which eliminates small pockets of gas from the thresholded image and thus prevents them from being deflated. We have found that such small pockets do not cause problems for registration and would introduce unnecessary deformations. The amount of erosion is two voxels, and of dilation, four. The extra voxels of dilation make the gas region in the binary image slightly larger than before, allowing the deflation to act on more of the intestinal wall than otherwise.

Using the refined thresholded image, we compute a deformation field just as for general deformable registration, by integrating an evolving velocity field $v(x, t)$ to get a deformation field $h_{\text{defl}}(x, t)$. In this case, the velocity field is computed using the equation

$$\nabla I(h_{\text{defl}}(x, t)) = Lv(x, t), \quad (5)$$

with L as in the equation for diffeomorphic registration (4). The only difference between (5) and (4) is that here the image force is simply given by the gradient of the image intensity. This causes the boundaries of the gas volumes to shrink towards the middle, as if deflating a balloon. We finally apply the resulting deformation field $h_{\text{defl}}(x)$ to the original image with gas.

Figure 4 shows an axial slice of a treatment image before and after gas deflation.

2.4. The Composite Transformation

We now describe how we combine the translation registration, the general fluid registration, and the gas deflation computation to calculate a single transformation from a planning image I_P to a treatment image I_T . We first perform the rigid translation to align the bones as well as possible. For this rigid registration, we choose an intensity window such that relatively dense bone appears white (maximum intensity), and other tissue appears black. We use a region of interest that includes the medial portion of the pelvis and excludes the femur outside the acetabulum. This computation gives us a translation vector τ .

We then apply the deflation algorithm to I_P and I_T to get two new images $I_{P\text{-defl}}$ and $I_{T\text{-defl}}$, with associated deformation fields $h_{P\text{-defl}}$ and $h_{T\text{-defl}}$ such that $I_{P\text{-defl}}(x) = I_P(h_{P\text{-defl}}(x))$, and similarly for $I_{T\text{-defl}}$. Finally, we apply deformable registration to $I_{P\text{-defl}}$ and $I_{T\text{-defl}}$, yielding the deformation field $h_{TP(\text{defl})}$. Then the full deformation field

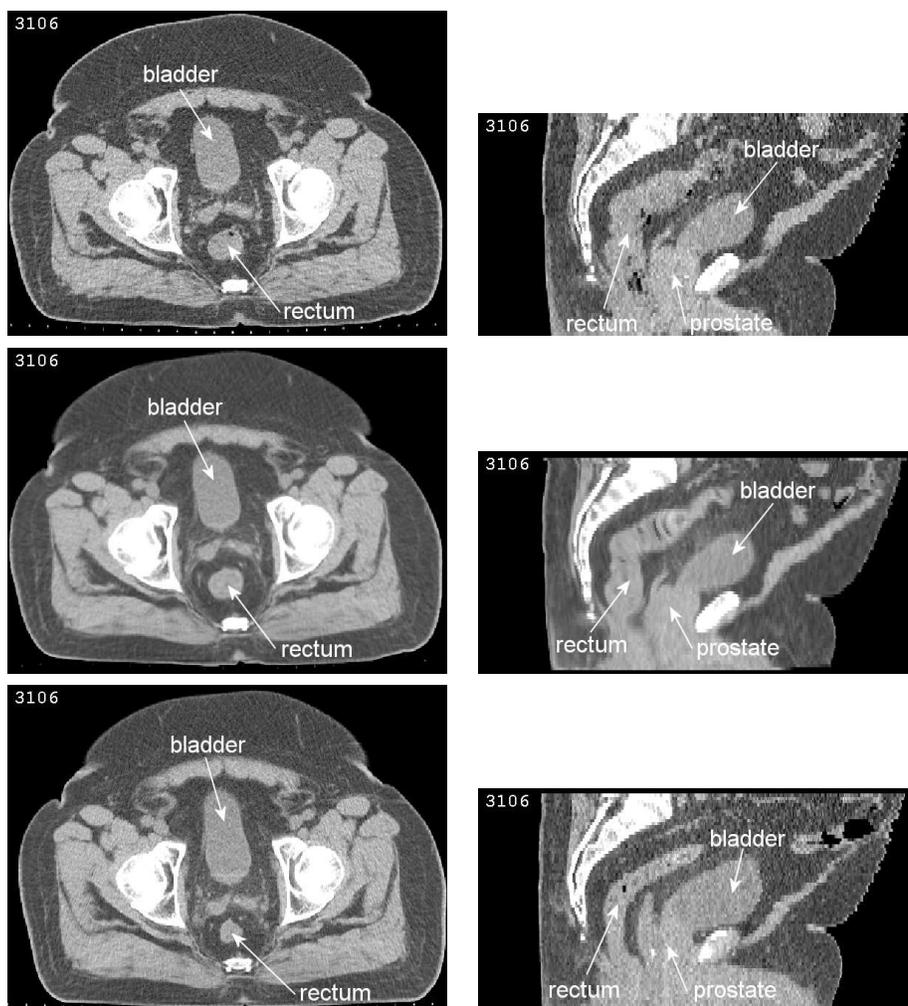


Figure 5. Example of deformable image registration. The first and last rows show axial and sagittal slices of the planning and treatment images. The second row shows the treatment image after deformable image registration, which brings the treatment image into alignment with the planning image. The improvement in soft tissue correspondence suggests that the registration procedure accurately captures internal organ motion. Note how the changes in size and shape of the bladder and rectum are accounted for.

warping I_T to the space of I_P is given by

$$h_{TP}(x) = h_{T\text{-defl}}(h_{TP(\text{defl})}(h_{P\text{-defl}}^{-1}(x))) + \tau.$$

Accordingly, the point x in the planning image corresponds to the point $h_{TP}(x)$ in the treatment image. This sequence of transformations can be represented as follows:

$$V_P \xrightarrow{h_{P\text{-defl}}^{-1}} V_{P\text{-defl}} \xrightarrow{h_{TP(\text{defl})}} V_{T\text{-defl}} \xrightarrow{h_{T\text{-defl}}} V_{T\text{-align}} \xrightarrow{+\tau} V_T.$$

Table 1. Parameters used in the regularizing operator $L = \alpha \nabla^2 + \beta \nabla \nabla \cdot + \gamma$, along with the maximum number of iterations permitted.

Scale	α	β	γ	Iterations
Coarse	0.01	0.01	0.001	150
Medium	0.01	0.01	0.001	75
Fine	0.02	0.02	0.0001	25
Deflation	0.02	0.02	0.0001	200

2.5. Multiscale Registration Implementation

For both rigid and deformable registration we use multiscale techniques to improve efficiency. We resample the images to 1/2 and 1/4 their original resolutions, and then apply our registration algorithm to the coarsest image first, using the result to initialize the algorithm on the next finer image. In the case of deformable registration, we interpolate the deformation field acquired at one resolution to generate the initialization for the immediately finer stage. The parameters we use for α , β , and γ in the definition of L depend on the coarseness of the scale. The values we have used are shown in table 1. For gas deflation we only do fine-scale calculations.

The runtime for the full registration algorithm is proportional to the size of the images being registered, and is dominated by the gas deflation and deformable registration computations, which require two 3D fast Fourier transforms (FFT) per iteration. Each FFT requires on the order of $n \log n$ floating point operations, where n is the number of voxels in each image. For our experiments, n ranges from 1,164,942 ($81 \times 102 \times 141$) to 7,912,905 ($187 \times 217 \times 195$), depending on the patient.

For deformable registration specifically, the time per iteration, averaged over all patients in our study, is 0.2 sec, 2.0 sec, and 22.7 sec for coarse, medium, and fine resolution computations, respectively. The average time for deformable registration was approximately 12.5 minutes per daily image. These results were obtained on a PC with 4GB of main memory and dual 3GHz Intel Xeon processors (although only one processor was used in the computations).

3. Automatic segmentation

The goal of image guidance in radiation therapy is to measure the changes over time of tumor and organs in both location and shape, so that the treatment can be adjusted accordingly. In our current ART practice we use manual contouring of organs for this purpose, but this is problematic because it is time consuming, and because there is considerable variation even when the same individual contours an image repeatedly on different days (Collier *et al* 2003). Instead, using image deformation, it is possible to carry the contours from the planning image to a daily image, deforming them to match the new image. This provides an automatic segmentation of the new image, based on the

manual segmentation of the planning image. In practice, the automatic segmentations must still be reviewed by a physician, but they need not be edited unless an error is found. In this section, we explain our method, and then present a statistical analysis of its accuracy and reliability.

The idea is to use the deformation fields to move the vertices of the contours from their locations in the planning image to the corresponding points in the treatment image (figure 6). This process does not result in a set of planar contours, since vertices will

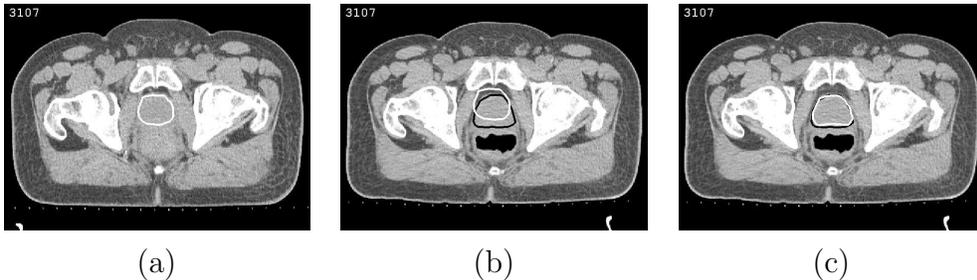


Figure 6. Example of automatic segmentation using deformable image registration. (a) Axial slice of a planning image with the prostate labeled by a white contour. (b) The same axial slice (in terms of planning coordinates) from a treatment image. The planned prostate position is shown in white, the actual prostate in black (both contours manual). (c) The same treatment image and manual (black) contour. The white contour is automatically generated by performing deformable image registration and applying the resulting deformation to the planning segmentation. The close agreement of the contours indicates that image registration accurately captures the prostate motion.

typically be moved out of plane to varying degrees. Therefore, instead of working with the contours directly, we first convert the sequence of contours to a surface model made up of triangles (figure 7) using an algorithm due to Amenta *et al* (2001). Then, we replace each vertex x in the model with $h(x)$, after which we slice the model with planes parallel to the xy axis to generate a new set of contours.

Figure 8 permits a visual assessment of the accuracy of our method. This figure is similar to figure 1 except that, instead of denoting the actual daily prostate positions, the contours represent the daily prostate positions deformed into the space of the planning image. Discrepancies between the deformed segmentations measure not only image registration uncertainty, but also intra-rater variability of the manual, treatment-day segmentations. In the rest of this section we quantify the accuracy of our segmentation method in more detail, with attention to human variability.

Our statistical analysis is based on comparing automatically generated segmentations to manual, hand-drawn segmentations. However, there is appreciable variation in manual segmentation, making it unreasonable to choose a particular manual segmentation as definitive. Groups have reported segmentation variation in a number of contexts, including brain tumors (Leunens *et al* 1993), lung cancer (Valley and Mirimanoff 1993, Ketting *et al* 1997), and prostate MR (Zou *et al* 2004). Rasch *et al* (1999) reported inter-user variabilities in the segmentation of the prostate in CT and MRI, finding over-

all observer variation of 3.5mm (1 standard deviation) at the apex of the prostate and an overall volume variation of up to 5% in CT.

Given this inter-rater variability, we assess our method by comparing our automatically generated segmentations with segmentations from manual raters. We then compare segmentations from different manual raters. We judge the accuracy and reliability of the automatic segmentations based on the standard of the measured inter-rater variability.

We have acquired CT scans for a total of 138 treatment days from 9 patients enrolled in our protocol. All of these images have been manually segmented by at least one expert. However, due to the time-consuming nature of manual segmentation, images from only 5 of these patients have been manually segmented by a second expert. We use the 65 images from these 5 patients for the analysis in this section. Eventually we plan to perform the same analysis for all of the patients enrolled in our protocol. Volume overlap statistics for the available segmented organs for all 9 patients are presented in Section 3.2.

The experimental setup is as follows. This study is based on a total of 65 CT images representing 65 treatment days for 5 patients. Each CT scan was collected prior to treatment on the Siemens Primatom scanner mentioned above, with a resolution of $0.098 \times 0.098 \times 0.3$ cm. Each planning image, as well as every treatment image, is manually segmented twice, once by rater *A* and once by rater *B*. For each patient, our method is used to compute the transformations h_i that deformably align the planning image with the treatment image for each treatment day i . Automatic segmentations are generated for each treatment image by applying h_i to a segmentation in the planning image. We consider our automatic method for producing segmentations as rater *C* (for “computer”). We use C_A and C_B to represent treatment image segmentations

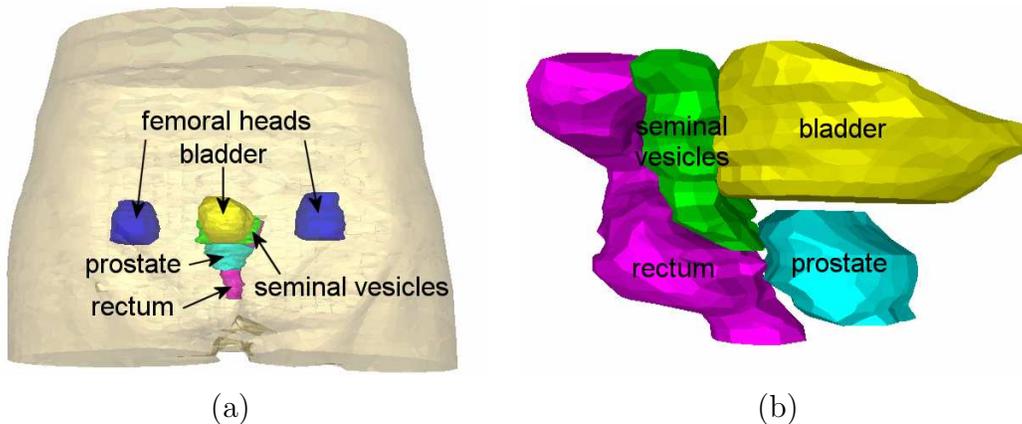


Figure 7. Visualization of organ segmentations. Panel (a) is an anterior view of a 3D rendering displaying segmentations of the skin, prostate, rectum, bladder, seminal vesicles, and femoral heads. Panel (b) shows a lateral view of the prostate, rectum, and bladder of the same patient. The surfaces are constructed by tiling manually drawn contours.

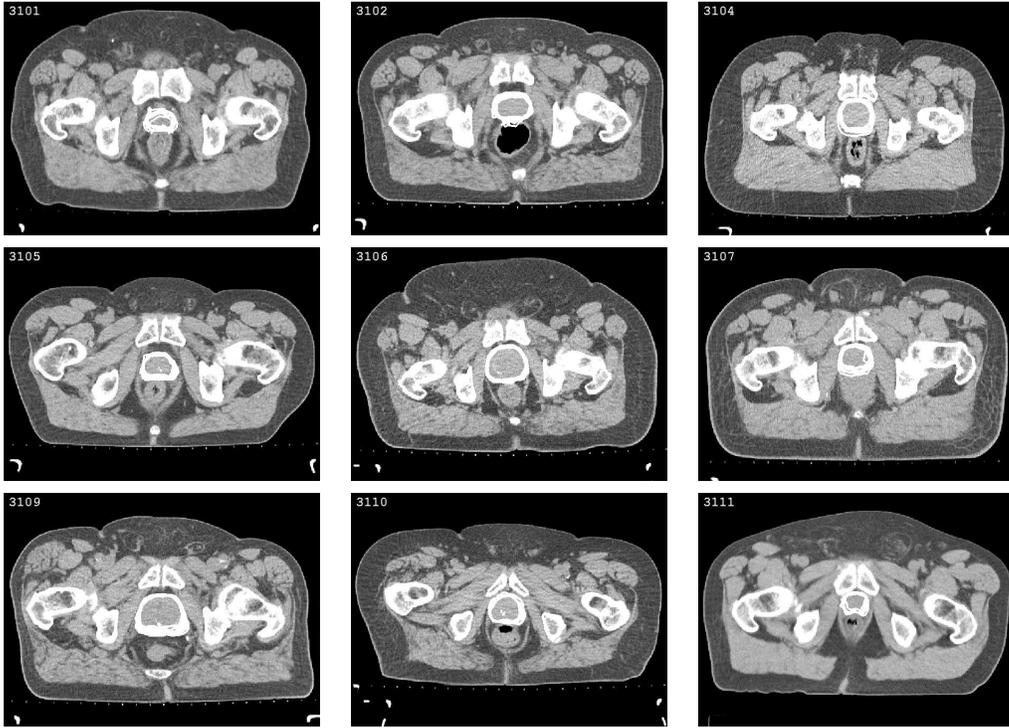


Figure 8. Visualization of the result of image registration algorithm. The images show manual segmentations of each daily image deformed into the space of the planning image. The close agreement of the deformed segmentations with the position of the prostate in the planning images provides evidence for the accuracy of the image registration algorithm along the prostate boundary.

that have been automatically generated by deforming the manual planning image segmentations drawn by raters A and B , respectively. Therefore, there are a total of 4 segmentations for each treatment image: 2 manual segmentations (A and B) and 2 automatic segmentations (C_A and C_B).

For each patient and for each treatment day, there are 6 pairwise comparisons that can be made from the set of 4 segmentations. We report data on 5 of these comparisons: AB , comparing manual segmentations by rater A against those by rater B ; C_AA and C_BB , comparing automatic segmentations with manual segmentations produced by the same rater; and C_AB and C_BA , comparing automatic segmentations with manual segmentations produced by a different rater. It should be emphasized that the automatic segmentations are produced by transforming manual planning segmentations produced by either rater A or rater B . Thus, we expect the same rater comparisons to be more favorable than the cross rater comparisons, which will be influenced by inter-rater variability.

In the rest of this section, we present the results of this experiment when measuring centroid differences and volume overlap of segmentations. We also show radial distance maps, which help us understand which regions of the prostate have the largest segmentation differences.

3.1. Centroid Analysis

The centroid of the prostate is especially important for radiation treatment planning and therapy because it is the origin, or isocenter, for the treatment plan. To measure the accuracy of our automatic segmentations with respect to centroid measurement, we compare the centroid of each automatic segmentation with the centroid of the corresponding manual segmentation.

First we consider the question: Are the centroids of the automatic segmentations systematically shifted with respect to the manual rater segmentations? Let S_A^i , S_B^i , $S_{C_A}^i$, and $S_{C_B}^i$ denote the prostate segmentations from raters A , B , C_A , and C_B , respectively, for image i . Let $\mathcal{C}(\cdot)$ be a function that returns the centroid (in \mathbb{R}^3) of a segmentation. In order to determine whether the centroids of the automatic segmentations are systematically shifted in any particular direction, we examine the distribution of the centroid differences $\mathcal{C}(S_{C_A}^i) - \mathcal{C}(S_A^i)$, $i \in 1, 2, \dots, N$ (and similarly for C_B). Likewise, to test for systematic shifts between manual raters A and B , we examine the distribution $\mathcal{C}(S_B^i) - \mathcal{C}(S_A^i)$. Figure 9 (a) shows box-and-whisker plots of these differences for the BA , $C_A A$, and $C_B B$ comparisons. The differences in the lateral

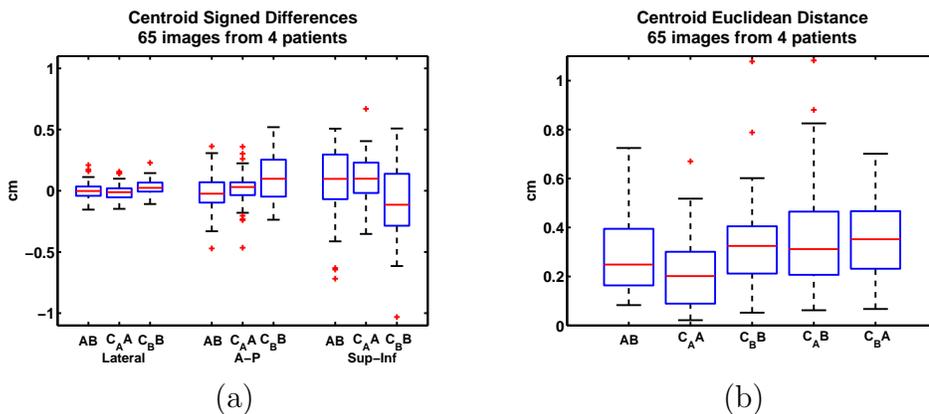


Figure 9. (a) Centroid differences measured in the lateral (X), anterior-posterior (Y), and superior-inferior (Z) directions. The horizontal lines on the box plots represent the lower quartile, median, and upper quartile values. The whiskers indicate the extent of the rest of the data, except that outliers, which fall more than 1.5 times the interquartile range past the ends of the box, are denoted with the ‘+’ symbol. (b) Euclidean distance between segmentation centroids.

(X), anterior-posterior (Y), and superior-inferior (Z) directions are measured separately. Summary statistics are provided in Table 2. It can be seen from this data that there is no significant shift between centroids of the computer generated segmentations and rater A’s manual segmentations in the lateral and A-P directions. There *is* a significant shift ($p < 0.001$ for two-tailed t-test) in the sup-inf direction of approximately 0.09 cm, which is less than one third of the sup-inf image resolution (0.3 cm). For the $C_B B$ comparisons we find significant shifts in the lateral and A-P directions of approximately 0.03 cm and 0.12 cm, respectively, which are at or less than the voxel resolution in these

Table 2. Summary statistics for centroid difference distributions. The mean, standard deviation, and 95% confidence interval for the mean are reported.

	Lateral (X)			A-P (Y)			Sup-Inf (Z)		
	BA	C_{AA}	C_{BB}	BA	C_{AA}	C_{BB}	BA	C_{AA}	C_{BB}
Mean	0.00	-0.01	0.03	-0.01	0.02	0.12	0.07	0.10	-0.07
STD	0.07	0.07	0.06	0.15	0.13	0.18	0.28	0.18	0.28
95% CI	-0.02	-0.03	0.01	-0.05	-0.02	0.07	0.00	0.05	-0.14
	0.02	0.00	0.04	0.02	0.05	0.16	0.14	0.14	0.00

Table 3. Summary statistics for centroid distance distributions.

	Euclidean Distance				
	AB	C_{AA}	C_{BB}	C_{AB}	C_{BA}
Mean	0.29	0.21	0.32	0.37	0.35
Median	0.25	0.20	0.32	0.31	0.35
Max	0.72	0.67	1.08	1.08	0.70
STD	0.16	0.13	0.17	0.22	0.15
IQR	0.23	0.21	0.19	0.26	0.24

dimensions. The comparison between manual raters shows that there is a significant shift in the Sup-Inf direction of approximately 0.07 cm.

In the lateral and Sup-Inf directions, the standard deviation of the manual AB comparisons is as large or larger than the standard deviation of the C_{AA} and C_{BB} comparisons. In the A-P direction, the standard deviation of the C_{BB} comparisons is slightly higher than the manual comparison.

Next we examine the Euclidean distance measured between segmentation centroids. Figure 9 (b) shows box-and-whisker plots of these distances. Summary statistics for these data are presented in table 3. As the distributions of these distances are not approximately normal, we report medians and interquartile ranges as well as means and standard deviations.

All of the mean distances are within image resolution. We tested for equality of the means of these distributions using paired t-tests. The C_{AA} mean distance is significantly less than the AB mean distance ($p < 0.001$) while there is no significant difference between the C_{BB} and AB mean distances. As expected, we see that centroids of the automatically generated segmentations are consistently closer to same-rater manual segmentations than cross-rater manual segmentations.

We conclude that the automatic segmentation method is comparable to human

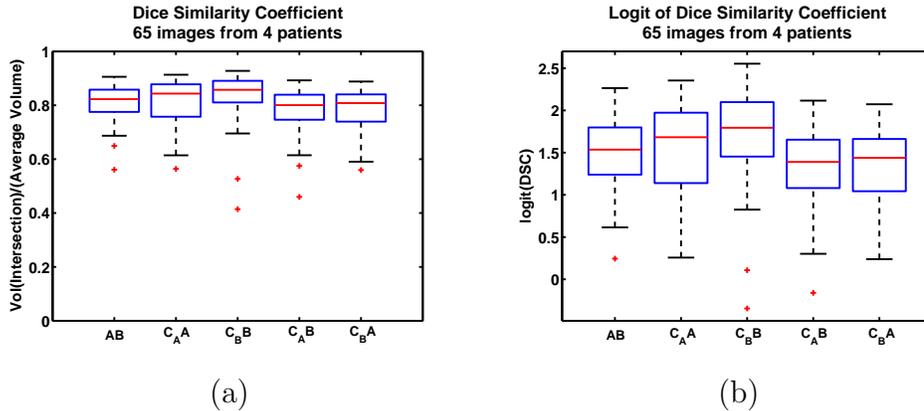


Figure 10. Dice Similarity Coefficient (DSC) and logit DSC.

raters in accuracy for estimating centroids and, as judged by the error bars and standard deviations, at least as reliable. However, there are outliers, with the maximum centroid distance being over 1 cm. For this reason, segmentations should be reviewed by a physician before being used in planning.

3.2. Volume Overlap Analysis

To measure the coincidence between volumetric segmentations of the prostate we use the Dice Similarity Coefficient (DSC) of Dice (1945). For two segmentations, S_1 and S_2 , the DSC is defined as the ratio of the volume of their intersection to their average volume:

$$\text{DSC}(S_1, S_2) = \frac{\text{Volume}(S_1 \cap S_2)}{\frac{1}{2}(\text{Volume}(S_1) + \text{Volume}(S_2))} \quad (6)$$

The DSC has a value of 1 for perfect agreement and 0 when there is no overlap. A DSC value of 0.7 or greater is generally considered to indicate a high level of coincidence between segmentations (Zijdenbos *et al* 1994, Zou *et al* 2004). The DSC can be derived from the kappa statistic for measuring chance-corrected agreement between independent raters (Zijdenbos *et al* 1994).

Figure 10 (a) shows a box-and-whisker plot of the Dice similarity coefficient for each comparison. The mean DSCs for the $C_A A$ and $C_B B$ comparisons are 0.82 (STD=0.08) and 0.84 (STD=0.08), respectively, indicating that the automatic segmentations have generally good coincidence with the manual segmentations. The mean DSC for the two manual raters was similar (mean=0.81, STD=0.06).

A similar study, carried out by Zou *et al* (2004), assessed the reliability of manual prostate segmentations in interoperative MR images. They report a mean DSC for manual raters of 0.838. Note that because prostate boundaries are more evident in MR images than in CT images, manual raters are likely to segment MR images more reliably than CT images.

Table 4. Summary statistics for the DSC measures.

	AB	C_{AA}	C_{BB}	C_{AB}	C_{BA}
Mean	0.81	0.82	0.84	0.78	0.78
Median	0.82	0.84	0.86	0.80	0.81
STD	0.06	0.08	0.08	0.08	0.08
IQR	0.08	0.12	0.08	0.09	0.10

Table 5. Comparison of automatic segmentation to manual segmenter A via the DSC and LDSC. This is the full set of segmenter-A segmentations that we have processed.

	DSC(S_1, S_2)		LDSC(S_1, S_2)	
	Prostate	Bladder	Prostate	Bladder
n	76	20	76	20
Mean	0.801	0.816	1.466	1.576
Median	0.825	0.826	1.554	1.557
STD	0.081	0.078	0.494	0.539
IQR	0.121	0.133	0.804	0.034

To evaluate the DSC distributions we use the logit of the DSC (LDSC), defined by

$$\text{LDSC}(S_1, S_2) = \ln \left(\frac{\text{DSC}(S_1, S_2)}{1 - \text{DSC}(S_1, S_2)} \right).$$

Agresti (1990) has shown that for large sample sizes (in the case of our prostate segmentations, the number of voxels is approximately 20 000), LDSC has a Gaussian distribution. Figure 10 (b) shows a box-and-whisker plot of the LDSC values for each comparison.

In order to test for a significant difference between the AB and C_{AA} or C_{BB} comparisons we performed paired t-tests on the LDSC values. A one-tailed test shows that the DSCs for the C_{BB} comparisons are significantly ($p < 0.001$) greater than the DSCs for the AB comparisons. We found no significant difference between the C_{AA} and AB comparisons ($p = 0.12$ for a two-tailed test). Therefore, the automatic segmentations coincide with the manual segmentations at least as well as a second manual rater.

Table 5 summarizes the manual-versus-automatic comparison for segmenter A only, for all patients that have been processed. After the first five treatment days, the bladder typically was not segmented.

3.3. Radial Distance Maps

Manual segmenters tend to find some portions of the prostate more difficult to segment than others. For instance, in CT there is often little or no apparent contrast between the prostate and bladder. Thus it makes sense to examine segmentation variability as a function of position on the prostate. For two segmentations X and Y of the same

image, we can visualize the deviation by choosing the centroid of X as a reference point, and considering, for each ray emanating from the centroid, the distance between the intersection points of the ray with X and Y . For each surface, we choose the first point that the given ray intersects that surface; typically there is only one. This procedure produces a distance for each radial direction, which can be plotted on the surface of a sphere, producing a *radial distance map*. This radial distance map is inspired by that of Mageras *et al* (2004), but we use a slightly different definition. To display the spherical map, we use the cartographic equal-area Mollweide projection. Since the patients are all scanned in a consistent orientation, different radial distance maps can be compared directly, and average maps can be computed point by point. Figure 11 shows the mean radial distance at each point for the cases analyzed in this section. Notice that the largest variation is generally found in the superior direction, which is consistent with the observed difficulty of detecting the boundary between prostate and bladder.

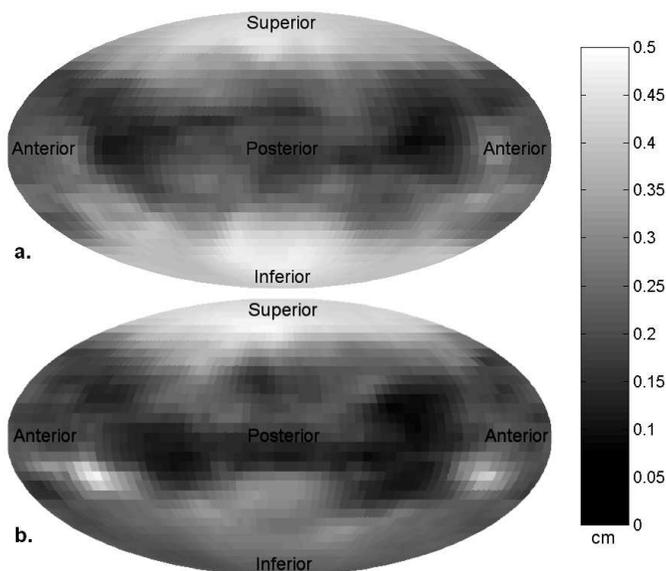


Figure 11. Radial distance maps, for the prostate. Map a: Mean radial distance between segmentations A and B (human raters). Map b: Mean distance between A and C_A segmentations (human and computer).

4. Dosimetric evaluation of image-guided radiotherapy

The day-to-day effects of organ motion and setup error can be illustrated by computing a dose volume histogram based on the observed organ location on each day. Figure 13 (a) shows DVHs for the first four days of treatment for patient 3102 of our protocol. The DVHs were computed by calculating the dose distribution based on the image for the given day, and applying that distribution to the organ segmentation computed by deforming the planning segmentations. Day 2 has a particularly severe cold spot, a fact confirmed by a comparison of the planning and treatment images. In figure 12, contours

for the prostate from both the planning day and treatment day 2 are shown, along with isodose lines for 95%, 98%, and 100% of prescribed dose.

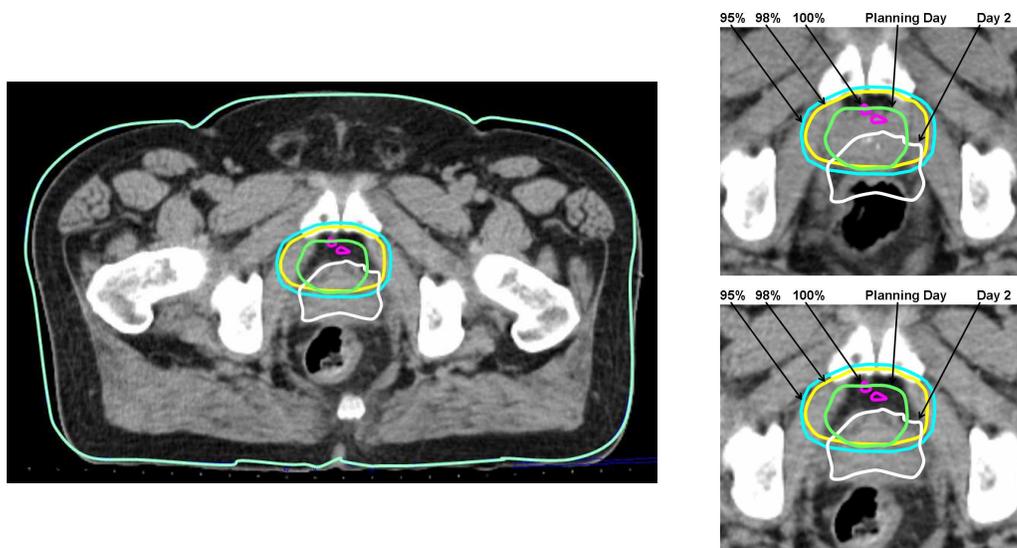


Figure 12. The position of the prostate in patient 3102 at day 2 compared to the time of planning. **Left:** The treatment image from day 2, shown with a skin contour from the planning day. **Top right:** Planning image. **Bottom right:** Day 2 image. On all three images, the location of the prostate is shown for both days, along with isodose curves at the 95%, 98%, and 100% level.

The top panel shows a full axial slice of the treatment image from Day 2, with an overlaid skin contour from the planning image as an indication of setup accuracy. The bottom two panels show closer views of the prostate from the planning and Day 2 images. In the slice shown, roughly half of the prostate appears to lie outside the 95% dose line.

It is not possible to directly combine a series of DVHs to produce an accurate DVH for the total dose delivered, because each DVH only indicates how great a volume from a given day received a specific dose. To combine information from different days, one needs to know the daily dose received by each voxel. Bortfeld *et al* (2004) provide a survey on the statistical effects of organ motion on dose distributions, using a rigid model. In the rest of this section we will describe how to assess total delivered dose in actual cases, considering deformation, by applying the displacement fields h computed from deformable image registration. Yan *et al* (1999), Birkner *et al* (2003) and Schaly *et al* (2004) have all described similar approaches, considering both raw and effective dose. However, their image registration algorithms require either that fiducial points be manually selected in the images, or that all of the images be segmented manually. In addition, none of these methods permit the range of deformations allowed by the fluid model.

4.1. Total Delivered Dose

Let the dose per fraction, as a function of position $x \in V$, be given by $D(x)$. Then the dose received at treatment i , by the tissue originally at x , is given by $D(h_i(x))$, and the total dose received by that tissue voxel over the course of treatment is given by

$$D_{\text{Tot}}(x) = \sum_i D(h_i(x)).$$

Using this formula, we can compute a distribution for total delivered dose, in the frame of reference of the planning day. Using the organ segmentations from the planning image, we can calculate DVHs that correctly reflect the variation in dose distribution over time. Figure 13 shows a series of delivered DVHs for increasing sets of treatment days. Before the histograms were computed, the dose distributions were normalized to the same prescription dose of 78 Gy. For instance, for the single-fraction DVH, the prescribed total dose was 2 Gy, so the dose to each voxel was multiplied by 39. As expected,

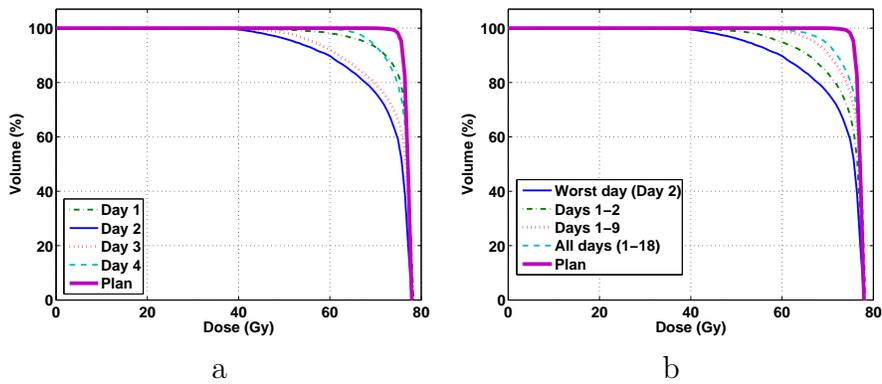


Figure 13. (a) Daily treatment prostate DVHs for each of the first four days, compared to the planned DVH. (b) Dose volume histograms for delivered dose, estimated over increasing sets of days. These are compared against the planned dose. All doses are normalized to a prescription dose of 78 Gy.

the quality of the DVH improves as the number of treatments being accumulated is increased, and we would expect further improvement given images from all 39 treatment days. But note that the DVH is still quite poor even based on 18 treatments, and that it only improved modestly over the 9-treatment DVH.

4.2. Effective Cumulative Dose

The difficulty with the measure D_{Tot} is that the biological effect does not depend simply on the total dose received, but also on the way it is distributed into fractions. Consider a volume of cells irradiated to a dose D over a time that is short relative to that required for cell repair to occur. Then the linear quadratic (LQ) model (Fowler 1989) gives the following estimate of the survival fraction SF of the cells in the volume:

$$\text{SF}(D) = e^{-\alpha D - \beta D^2}$$

Now let $T = (D_1, D_2, \dots, D_N)$ be a series of *varying* doses separated by time for cell recovery. In our situation, the relevant volume of tissue is a voxel x and, for each i , $D_i = D(h_i(x))$. Assuming that cell proliferation is negligible, the survival fraction for the treatment T will be given by

$$\text{SF}(T) = \prod_i e^{-\alpha D_i - \beta D_i^2} = \exp\left(\sum_i -\alpha D_i - \beta D_i^2\right).$$

Just as with uniform fractionation, one can construct the Biological Effective Dose (BED; Fowler 1989, Barendsen 1982). The BED is the dose that, if delivered in a series of fractions so small that the β term may be ignored, would kill the same number of cells as the actual dose in question. That is, we define $\text{SF}(\text{BED}) = e^{-\alpha \cdot \text{BED}}$, and compute the BED for a particular treatment regimen T by setting $\text{SF}(\text{BED}) = \text{SF}(T)$ and solving to obtain

$$\text{BED}(T) = \sum_i D_i + \frac{D_i^2}{\alpha/\beta}.$$

Then, following the analysis for the total delivered dose, we can define the total BED for a tissue voxel x as follows (see also Yan *et al* 1999, Birkner *et al* 2003, Schaly *et al* 2004):

$$\text{BED}_{\text{Tot}}(x) = \sum_i D(h_i(x)) + \frac{D(h_i(x))^2}{\alpha/\beta}. \quad (7)$$

To illustrate, figure 14 makes two comparisons. The planned BED is compared to the delivered BED, to indicate the differences due to organ motion. Also, the planned total dose is shown, to indicate the significance of the biological effect. For the purposes of illustration we assumed an α/β value of 1.5 Gy, which is at the low end of current estimates (Fowler *et al* 2001). The large difference between the planned BED and planned total dose reflects an assumption, embodied in the low value chosen for α/β , that prostate cancer is highly sensitive to the per-fraction dose. Larger values of α/β would bring the BED curves closer to the total planned dose curve. The delivered BED was estimated based on the 18 treatment images. That is, the delivered BED was computed by applying Equation 7 to the appropriate 18 dose distributions and deformation fields, with each distribution based on a prescription dose of 2 Gy/f. The resulting distribution represented the biological effect of the 18 treatments for which image data was available, so that the prescribed dose level for those prescriptions was 36 Gy. The resulting distribution was then normalized to a 78 Gy prescription dose by applying a scale factor of 78/36. As with raw dose accumulation, this estimate does not account for the improvement in the distribution that would result from averaging together a greater number of random motions.

Because of evidence indicating that prostate tumors may have α/β values comparable to healthy tissue, there is now considerable discussion of hypofractionation for prostate cancer (Kupelian *et al* 2002, Brenner 2003, Craig *et al* 2003). Figure 15 shows DVHs of accumulated BED for four values of α/β , assuming a regimen of 50 Gy delivered over 16 fractions, similar to one described by Logue *et al* (2001).

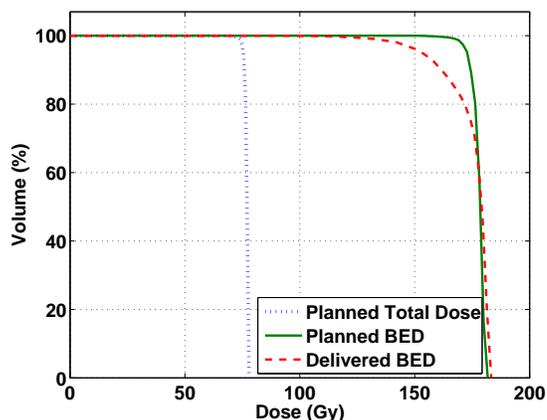


Figure 14. DVHs for Planned total dose, planned BED, and delivered BED. Delivered BED is modeled from a sample of 18 out of 39 treatment days. $\alpha/\beta = 1.5$.

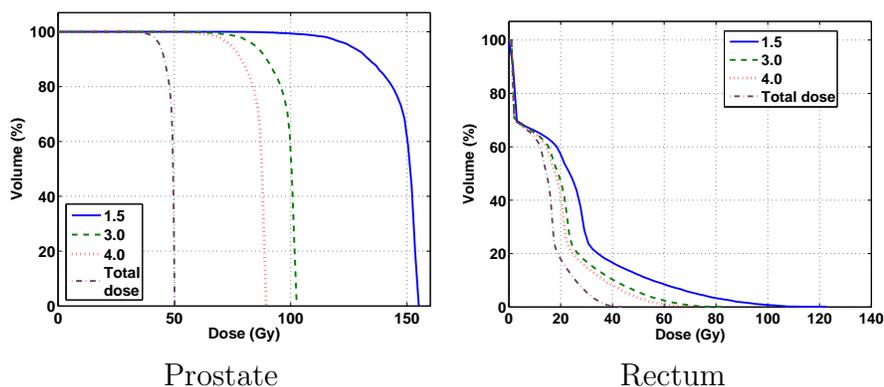


Figure 15. Biologically effective DVHs, assuming four possible values of α/β , and 50 Gy delivered in 15 equal fractions.

5. Discussion

We have described how large deformation image registration can be used for automatic segmentation and dose accumulation in the course of image guided radiation therapy. Our image registration technique is fully automatic, permits large deformations, and ensures a smooth one-to-one correspondence between two images. We use a variation of the registration method to eliminate bowel gas when it occurs, so that images can be brought into a meaningful correspondence.

For segmentation purposes, we compute the deformation that transforms the planning image to match the daily treatment image, and apply that deformation to the initial manual contours. We have validated this method by comparing the automatic segmentations to manual segmentations produced by the same segmenter who generated the original planning segmentations. Based on centroid difference and the DSC measure of volume overlap, we find that the automatic deformations of a planning segmenter correspond at least as closely to the daily segmentations of the same segmenter, as do

daily segmentations by a different individual. Although, in clinical practice, it will be necessary for a physician to check the segmentations, our data indicate that the number requiring modification will be small.

We also show how to use our registration method to estimate the amount of dose delivered to the patient over time, as a function of position within the imaged area. In a single case study we compare daily DVHs to both the planned DVH and to cumulative DVHs, observing that, as expected, the accumulation of multiple fractions tends to improve the correspondence between delivered and planned DVH, though we still find a pronounced difference based on 19 images. We also consider the accumulation of biologically effective dose. For 39 fractions, accumulated BED is very close to accumulated dose, but hypofractionation schemes lead to a greater difference. In the future we intend to apply these dose accumulation measures to assess the effectiveness of protocols both planned and currently in use in our clinic.

Acknowledgments

We thank Gregg Tracton for his assistance in organizing and processing the CT data sets, and Joshua Stough for his volume computation software. We also thank Gig Mageras, Michael Lovelock, and Paul Keall for helpful discussions. This work was supported by the DOD Prostate Cancer Research Program DAMD17-03-1-0134.

References

- Agresti A *Categorical Data Analysis* Wiley Interscience 1990
- Amenta N, Choi S and Kolluri R K The power crust In *ACM Symposium on Solid Modeling and Applications* pages 249–260 2001
- Barendsen G W 1982 Dose fractionation, dose rate and iso-effect relationships for normal tissue responses *International Journal of Radiation Oncology*Biology*Physics* **8** 1981–1997
- Beg M F, Miller M I, Trounevè A and Younes L 2005 Computing large deformation metric mappings via geodesic flows of diffeomorphisms *International Journal of Computer Vision* **61** 139–157
- Birkner M, Yan D, Alber M, Liang J and Nüsslin F 2003 Adapting inverse planning to patient and organ geometrical variation: Algorithmn and implementation *Medical Physics* **30** 2822–2831
- Bookstein F L 1989 Principal warps: Thin-plate splines and the decomposition of deformations *IEEE Transactions on Pattern Analysis and Machine Intelligence* **11** 567–585
- Booth J T *Modelling the Impact of Treatment Uncertainties in Radiotherapy* PhD thesis Adelaide University 2002

- Booth J T and Zavgorodni S F 2001 The effects of radiotherapy treatment uncertainties on the delivered dose distribution and tumour control probability *Australasian Physical & Engineering Sciences in Medicine* **24** 71–78
- Bortfeld T, Jiang S B and Rietzel E 2004 Effects of motion on the total dose distribution *Seminars in Radiation Oncology* **14** 41–51
- Brenner D J 2003 Hypofractionation for prostate cancer radiotherapy—what are the issues? *International Journal of Radiation Oncology*Biography*Physics* **57** 912–914
- Christensen G E, Carlson B, Chao K S C, Yin P, Grigsby P W, N K, Dempsey J F, Lerma F A, Bae K T, Vannier M W and Williamson J F 2001 Image-based dose planning of intracavitary brachytherapy: registration of serial-imaging studies using deformable anatomic templates *International Journal of Radiation Oncology*Biography*Physics* **51** 227–243
- Christensen G E, Joshi S C and Miller M I 1997 Volumetric transformation of brain anatomy *IEEE Transactions on Medical Imaging* **16** 864–877
- Christensen G E, Rabbitt R D and Miller M I 1996 Deformable templates using large deformation kinematics *IEEE Transactions On Image Processing* **5** 1435–1447
- Collier D C, Burnett S S C, Amin M, Bilton S, Brooks C, Ryan A, Roniger D, Tran D and Starkschall G 2003 Assessment of consistency in contouring of normal-tissue anatomic structures *Journal of Applied Clinical Medical Physics* **4** 17–24
- Craig T, Moiseenko V, Battista J and Van Dyk J 2003 The impact of geometric uncertainty on hypofractionated external beam radiation therapy of prostate cancer *International Journal of Radiation Oncology*Biography*Physics* **57** 833–842
- Csernansky J G, Joshi S, Wang L, Haller J W, Gado M, Miller J P, Grenander U and Miller M I 1998 Hippocampal morphometry in schizophrenia by high dimensional brain mapping *Proceedings of the National Academy of Sciences* **95** 11406–11411
- Davatzikos C 1996 Spatial normalization of 3d brain images using deformable models *Journal of Computer Assisted Tomography* **20** 656–665
- Dice L R 1945 Measures of the amount of ecologic association between species *Ecology* **26** 297–302
- Fowler J, Chappell R and Ritter M 2001 Is α/β for prostate tumors really low? *International Journal of Radiation Oncology*Biography*Physics* **50** 1021–1031
- Fowler J F 1989 The linear quadratic formula and progress in fractionated radiotherapy *British Journal of Radiology* **62** 679–694
- Goitein M 1985 Calculation of the uncertainty in the dose delivered during radiation therapy *Math. Phys.* **5** 608–612
- Goitein M 1986 Causes and consequences of inhomogeneous dose distributions in radiation therapy *International Journal of Radiation Oncology*Biography*Physics* **12** 701–704
- Goitein M and Busse J 1975 Immobilization error: Some theoretical considerations *Radiology* **117** 407–412

- Grenander U and Miller M I 1998 Computational anatomy: An emerging discipline *Quarterly of Applied Mathematics* **56** 617–694
- Happersett L, Mageras G S, Zelefsky M J, Burman C M, Leibel S A, Chui C, Fuks Z, Bull S, Ling C C and Kutcher G J 2003 A study of the effects of internal organ motion on dose escalation in conformal prostate treatments *Radiotherapy and Oncology* **66** 263–270
- Horn R A and Johnson C R *Matrix Analysis* Cambridge University Press 1990
- Joshi S, Lorenzen P, Gerig G and Bullitt E 2003 Structural and radiometric asymmetry in brain images *Medical Image Analysis* **7** 155–170
- Joshi S, Miller M and Grenander U 1997 On the geometry and shape of brain sub-manifolds *Int. Journal of Pattern Recognition and AI* **11** 1317–1342
- Joshi S C and Miller M I 2000 Landmark matching via large deformation diffeomorphisms *IEEE Transactions On Image Processing* **9** 1357–1370
- Ketting C H, Austin-Seymour M, Kalet I, Unger J, Hummel S and Jacky J 1997 Consistency of three-dimensional planning target volumes across physicians and institutions *International Journal of Radiation Oncology*Biography*Physics* **37** 445–453
- Kupelian P A, Reddy C A, Carlson T P, Altsman K A and Willoughby T R 2002 Preliminary observations on biochemical relapse-free survival rates after short-course intensity-modulated radiotherapy (70 Gy at 2.5 Gy/fraction) for localized prostate cancer *International Journal of Radiation Oncology*Biography*Physics* **53** 904–912
- Leunens G, Menten J, Weltens C, Verstraete J and van der Scheuren E 1993 Quality assessment of medical decision making in radiation oncology: Variability in target volume delineation for brain tumours *Radiotherapy and Oncology* **29** 169–175
- Logue J P, Cowan R A and Hendry J H 2001 Hypofractionation for prostate cancer *International Journal of Radiation Oncology*Biography*Physics* **49** 1522–1522
- Lu W, Chen M, Olivera G H, Ruchala K J and Mackie T R 2004 Fast free-form deformable registration via calculus of variations *Physics in Medicine and Biology* **49** 3067–3087
- Mageras G S, Joshi S, Davis B, Pevsner A, Hertanto A, Yorke E, Rosenzweig K and Ling C C Evaluation of an automated deformable matching method for quantifying lung tumor motion in respiration-correlated CT images In *ICCR 2004*
- Miller M I, Trounev A and Younes L 2002 On the metrics and euler-lagrange equations of computational anatomy *Annual Review of Biomedical Engineering* **4** 375–405
- Rasch C, Barillot I, Remeijer P, Touw A, van Herk M and Lebesque J V 1999 Definition of the prostate in CT and MRI: a multi-observer study *International Journal of Radiation Oncology*Biography*Physics* **43** 57–66
- Schaly B, Kempe J A, Bauman G S, Battista J J and Dyk J V 2004 Tracking the dose distribution in radiation therapy by accounting for variable anatomy *Physics in Medicine and Biology* **49** 791–805

- Thirion J P 1998 Image matching as a diffusion process: an analogy with maxwell's demons *Medical Image Analysis* **2** 243–260
- Thompson P M and Toga A W 2002 A framework for computational anatomy *Computing and Visualization in Science* **5** 13–34
- Unkelbach J and Oelfke U 2004 Inclusion of organ movements in IMRT treatment planning via inverse planning based on probability distributions *Physics in Medicine and Biology* **49** 4005–4029 URL <http://stacks.iop.org/0031-9155/49/4005>
- Valley J F and Mirimanoff R O 1993 Comparison of treatment techniques for lung cancer *Radiotherapy and Oncology* **28** 168–173
- Wang H, Dong L, O'Daniel J, Mohan R, Garden A S, Ang K K, Kuban D A, Bonnen M, Chang J Y and Cheung R 2005 Validation of an accelerated 'demons' algorithm for deformable image registration in radiation therapy *Physics in Medicine and Biology* **50** 2887–2905
- Yan D, Jaffray D A and Wong J W 1999 A model to accumulate fractionated dose in a deforming organ *International Journal of Radiation Oncology*Biography*Physics* **44** 665–675
- Yan D, Lockman D, Brabbins D, Tyburski L and Martinez A 2000 An off-line strategy for constructing a patient-specific planning target volume in adaptive treatment process for prostate cancer *International Journal of Radiation Oncology*Biography*Physics* **48** 289–302
- Yan D, Vicini F, Wong J and Martinez A 1997 Adaptive radiation therapy *Physics in Medicine and Biology* **42** 123–132
- Zijdenbos A P, Dawant B M, Margolin R A and Palmer A C 1994 Morphometric analysis of white matter lesions in mr images: Method and validation *IEEE Transactions on Medical Imaging* **13** 716–724
- Zou K H, Warfield S K, Baharatha A, Tempany C, Kaus M R, Haker S J, Wells W M, Jolesz F A and Kikinis R 2004 Statistical validation of image segmentation quality based on a spatial overlap index *Academic Radiology* **11** 178–189